

METHOD AND DISPLAY FOR MULTIVARIATE CLASSIFICATION

Related Applications

ms 44
4/8/04
This application claims ^{benefit of} priority under 35 U.S.C. §119 from U.S. provisional
5 application serial number 60/240,403, filed October 13, 2000, and from U.S. provisional
application serial number 60/252,290, filed November 21, 2000.

Field of the Invention

10 The invention relates to methods of classifying tissues based on multivariate data for
diagnostic, prognostic or therapeutic purposes. The invention also relates to methods and
user interfaces for displaying such multivariate data to facilitate rapid decision making.

Background of the Invention

15 Classification of tissues historically depended on an examination of the gross
morphology and histology of the tissues. The inadequacy of these methods for classification
of tissues that are similar in appearance, such as different tumors arising from the same tissue
or organ, is known. More recently, methods of tissue analysis and classification have been
developed that rely on genetic analysis of tissues. While these genetic analytical methods are
more powerful for distinguishing tissue types, and are simple to practice, the methods also
20 generated quantities of data that are orders of magnitude greater than classical histology
methods. For example, gene expression profiling of tissues by the application of microarray
technology determines simultaneously the expression of thousands of genes. The challenge
of identifying the data useful in tissue classification from the raw data is a problem that
requires a solution that can be applied in clinical settings. Present statistical and data display
25 methods are not satisfactory for this application.

Thus it is of key importance in analytical processes for classification of unknown
tissues reduce large quantities of multivariate data to a quantity that can be readily analyzed.
A typical treatment of multivariate data is to generate binary statistics known to those in the
art. For example, a binary regression of two data sets associated with two objects will
30 produce a statistical relationship between the two objects within a statistical confidence.

Prior art data analysis tools have focused on describing objects of diagnostic
relevance to exhaustion, which in turn relied on using binary statistical methods to "collapse"
large volumes of data for decision making. These methodologies suffer from a requirement